

DOCUMENT RESUME

ED 324 775

EA 022 298

AUTHOR Dickson, LouAnn
TITLE Student Achievement and Career Ladder Status.
PUB DATE Apr 90
NOTE 24p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Career Ladders; Class Size; Elementary Secondary Education; *Incentives; Professional Development; *Program Evaluation; Scores; *Teacher Salaries
IDENTIFIERS *Mesa Public Schools AZ

ABSTRACT

The Career Ladder Teacher Incentive Program was designed by the Arizona legislature to attract, motivate, and retain talented teachers with both professional recognition and monetary incentives. The program was also intended to improve student achievement. Using carefully matched samples of career-ladder and non-career-ladder teachers in Mesa Public Schools, the study summarized in this report compared mean class scores between and within these groups on the Iowa Test of Basic Skills in a 2-year investigation. When these means were analyzed they did not achieve statistical significance in most cases. However, classes with career-ladder teachers scored an average of from 2 to nearly 6 points higher than those taught by non-career-ladder teachers. According to research (by Glass and Smith), as class size decreases, student achievement increases. Increasing student achievement by 6 percentiles in a class of 26.7 students (the Mesa District average) places these students' achievement levels near those expected of a class of 15 students. These data reflect only the first 2 years of career ladder implementation. Results will become clearer as goals and objectives are clarified, procedures and reporting become standardized, and wider teacher acceptance and participation become a reality. So far, the pilot program is achieving its goal. (13 references) (MLH)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED324775

STUDENT ACHIEVEMENT AND CAREER LADDER STATUS

BY
LOUANN DICKSON, PH.D.

MESA PUBLIC SCHOOLS
RESEARCH AND EVALUATION
549 NORTH STAPLEY DRIVE
MESA, ARIZONA 85203

EA 022 298

Paper presented at the Annual Meeting of the
American Educational Research Association,
Boston, MA, April 16-20, 1990.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

L. Dickson

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

STUDENT ACHIEVEMENT AND CAREER LADDER STATUS

INTRODUCTION

School boards, administrators, and teachers should cooperate to develop career ladders for teachers that distinguish among the beginning instructor, the experienced teacher, and the master teacher.

A Nation at Risk (pg. 31)

Interest in incentive plans for paying teachers can be traced to as early as the beginning of this century with the first recorded merit pay plan being established in Newton, Massachusetts in 1908. Interest continued in these plans, peaking after World War I and finally declining in the 1930's as single salary schedules were established for teachers based primarily on education and longevity. With the exception of a short and limited revival of merit pay issues in the 1950's, the concept of incentives continued to languish and decline (Robinson, 1984).

In the late 70's, a study found only 115 school districts nationwide using any kind of incentive pay for teachers: 239 districts had tried and subsequently abandoned incentive plans (ERS, 1978). By 1983, only 54 districts reported use or plans to use merit pay plans (ERS, 1983). The primary reason cited for failure of these plans was dissatisfaction with evaluation procedures (Robinson, 1984). Conventional wisdom had come to accept the idea that the only fair way to reward teachers was on the basis of educational attainment and longevity of service, that teaching productivity cannot be measured objectively, that student achievement as measured by test results cannot measure teacher performance and that incentive programs for teachers are unworkable because of teacher opposition (Robinson, 1984). For the middle part of the century, these assumptions virtually stalled all serious exploration of merit pay systems for teachers.

Publication of *A Nation at Risk* (USDE, 1983) riveted national attention on the dual problems of declining student achievement and the growing shortage of

competent teachers. Districts began to seek avenues to improve student learning and to attract academically talented students to teaching careers. Suddenly, approaches rewarding excellence in teaching with increased stature and pay began to appear as contingencies for legislative funding of educational programs at both the state and federal levels. Officials began insisting that increased funding for education be at least partly contingent upon merit or incentive pay for teachers. These legislative mandates added new impetus to the exploration of approaches to merit pay which had languished in obscurity for decades, primarily due to problems in finding ways to evaluate teachers.

Evaluation within these incentive programs has centered almost exclusively on teacher variables (Mitchell & Peters, 1988; Peterson, 1987). Within these teacher variables, focuses have been on both intrinsic and extrinsic reward systems at individual, group and organizational levels. These teacher evaluations employ such avenues as administrative reports, pupil reports, parent surveys and reports, teacher tests, professional activities, peer assessment, peer counseling, experience, job-related education and advanced education. Career ladder programs are one means of incorporating these various techniques into teacher evaluations. Variants of the career ladder approach are numerous but typically teachers are evaluated against increasingly stringent criteria as they move up the career ladder. Pay incentives are criteria-based at each level. Career ladder programs have been adopted with varying success in school districts in such states as Utah, Tennessee, and Texas. Although program formats and evaluation techniques vary widely, the goals of nearly all are to establish systems of accountability which reward excellence.

While concentrating on teacher variables, most of the approaches to evaluation have ignored the student and the effect of the programs on student achievement. Tennessee, which instituted a career ladder program as early as 1984, is just now beginning to examine ways to include student achievement and attitudes in the evaluation of teachers in that program (Cornett, 1987), but the primary emphasis still is on teacher variables. Only the state of Arizona to date, has mandated that evidence of benefit to student achievement from the program be demonstrated.

The Arizona legislature enacted a 5 year career ladder program (Career Ladder Teacher Incentive Program) and appointed a Joint Legislative Committee to approve and monitor pilot programs to be established in 15 school districts throughout the state. Cognizant of the diversity of districts within the state, each pilot program was to be locally designed and directed, but the legislation mandated that programs meet certain criteria. First, district teachers must be involved in development of the program. Second, improved student achievement must be the primary goal. Third, advancement on the ladder must reflect improved or advanced teaching skills and/or responsibilities. Advancement was not to be determined by either longevity or increased education.

Career ladder programs in Arizona currently involve approximately 10,000 educators, their students, and 12 million dollars in incentive funds (Packard & Dereshiwsky, 1988). Participating districts involved vary in size, locale and clientele with 2 of the 15 participating districts located within the Navajo Nation. Seven districts began programs in 1985-86, three more initiated programs in 1986-87, and five were instituted 1987-88. A summative evaluation of these programs was presented to the Legislature in the Fall of 1989.

Mesa Public School District, the largest in Arizona, initially established a committee to make recommendations and outline a possible career ladder program for submission to the legislature. A program was designed and submitted for teacher approval but was vigorously opposed and subsequently rejected by district teachers. A year later, a 10 member committee, comprised of seven educators recommended by the local teachers' association and three principals presented a proposal which was later adopted incorporating the criteria defined by the state and two additional guidelines:

1. Selection of Career Ladder teachers should be based on clear, predetermined criteria and not limited by quotas;
2. Educators should be able to choose whether or not to apply for career ladder and to withdraw from the plan at any time.

PROGRAM DESCRIPTION

Mesa Public Schools (MPS) initiated their career ladder program (CL) during 1986-87 (Mesa Public Schools, 1986). The program has two levels: Resident and Career. The Resident Level has two steps and the Career Level has three. For each level and each step, specific qualifications are defined, based on the teachers satisfactory demonstration of competence in three primary educator roles: instruction, management and counseling. Within each of these primary roles the teacher must demonstrate ability to plan, implement, and evaluate. At higher CL levels teachers must also demonstrate advanced skills in adaptation, integration and the ability to innovate and lead.

Each of these primary roles is defined by a number of specific skills. For example, within the broad role of management, the teacher's ability to plan would be demonstrated by such skills as formulating a standard of behavior and developing a system for keeping records. The principal or immediate supervisor evaluates each applicant yearly for the Career Ladder using the basic and advanced educator skills definitions. The number of skills needed for placement at each level increases as the teacher moves up the ladder.

A second essential component of the Career Ladder is the verification of student achievement. Each fall applicants submit a plan that specifies their objectives for student learning, how they will assess this and their goals for each student. A trained team of teachers evaluates each plan which may be returned to the teacher for revision if the objectives, assessment or goals are not clear or appropriate. At the end of the school year, each teacher submits a report and evaluation of their results based upon their students' actual performance.

Each step of Mesa's Career Ladder Plan has a specified salary. Teachers who qualify for a higher salary than that provided by the district's standard schedule receive an addendum to their contract for the difference. As district salary levels increase, limitations on the amount of the addendum are encountered. This approach is particularly rewarding for those in the first few years of teaching since they can receive professional salaries early in their careers, but problematic for teachers who have been in the district

longer periods of time (Dickson & Guy, 1990). Although analyses of these pay schedules demonstrate substantial benefits even at the highest district level, there has been acrimony among more experienced teachers from this perceived inequity, and their initial participation rates were low.

In addition to the basic CL steps, Career Ladder educators, may apply for paid positions that require them to spend time and energy beyond their contractual responsibilities to the district. Positions available include demonstration teachers, elementary instructional improvement coaches and mentoring. In addition, positions directly related to implementation of the Career Ladder program itself are available to Career Ladder teachers and include such responsibilities as school liaisons and evaluators of portfolios, student achievement plans and optional extender proposals. Career Ladder teachers may also apply for funds to design and implement special programs which meet a variety of needs.

Mesa Public Schools Career Ladder program is now completing its third year of implementation. The first year, 14% of district teachers were placed on Career Ladder status and the second year 17% were placed. At the beginning of the third year, 30% of the district's teachers have applied, indicating growing interest in the program.

When the Arizona legislature mandated a career ladder program to attract, retain and reward talented teachers, they automatically built the student achievement component into the program. It seems safe to assume that a talented teacher will motivate and instruct students in such a way as to increase their achievement over that of, say, a mediocre teacher. In a program designed to select talented and master teachers from the overall teacher pool, it seems reasonable to expect that students of these talented teachers will be performing at a higher level than students of teachers from the overall pool. If this is *not* true, one might question the efficacy of the selection process and wonder what kind of teachers, in fact, are selected by the criteria established. Certainly that would not appear to meet the spirit of either the national reform movement or the legislative mandate.

If we analyze career ladder programs, we find two basic components: a *selection process* in which teachers are asked to demonstrate that they meet the criteria for inclusion, and a *retention process* in which teachers demonstrate continuing professional development in order to meet ever more stringent steps on the ladder. As we have argued, the selection process should select teachers whose students excel, but it also seems reasonable to assume that the retention process will provide these talented teachers with an impetus for professional development that will further increase their efficacy in the instructional process and subsequently further increase achievement of their students. As Career Ladder teachers define objectives and evaluate various approaches to instruction to provide necessary lines of evidence for retention in the program, resultant professional growth will manifest itself in improved student achievement scores over time. We believe that if standardized scores for students of Mesa's Career Ladder teachers are examined they would demonstrate this initial elevation and, over time, present a pattern of further increase.

Several problems arise with this kind of evaluation research because the MPS program is still in a relatively early stage of development. Test scores are only available at this time for the first 2 years of the program. Examination of testing results from the first year will give us a picture of how well the selection process is working and scores from subsequent years will provide us with information on professional developments and its impact on the student.

Since the program is relatively new, it is conceivable that for whatever reasons, talented teachers who eventually might be included in the career ladder group have not yet applied for career ladder status. If, as we believe, student achievement scores for these teachers are higher than those of teachers who would not be selected for career ladder status, scores for the non-career ladder group would be inflated. As these teachers are absorbed into the career ladder program, the conundrum will resolve and increased separation of scores should become evident.

Another consideration that must be addressed in this context is that change does not occur in a vacuum. As Career Ladder teachers develop programs, explore new instructional techniques and report successes, sharing of these

approaches and techniques can provide improved role models and collegial consultation that will ultimately improve instructional quality of non-career ladder teachers also. If this occurs, we may expect an initial dip in standardized testing scores for the overall teacher pool as Career Ladder teachers are selected out, with a subsequent increase as successful approaches to instruction are shared among colleagues.

Given these caveats, we designed a study to track trends in achievement scores. If the program at MPS is selecting talented teachers for monetary incentives, we predict a consistent pattern across grades and course content of higher achievement scores for Career Ladder teachers over non-career ladder teachers and for higher level Career Ladder teachers over lower level Career Ladder teachers.

METHODOLOGY

SUBJECTS

Year One. Our sample consisted of 224 teachers of grades 1-6 in the Mesa Public Schools of whom 112 were currently in the Career Ladder program (CL) and 112 were non-career ladder teachers (NCL). The two groups were matched on the basis of class regrouping, school characteristics, grade, and class size. Information on teachers experience was included for statistical control although sample size precluded matching on this variable.

Initial inclusion in either group was determined on the basis of whether reading, math and language classes regrouped within schools. Regrouping occurs when students are reassigned for a particular course content to a different teacher. When regrouping occurs, we are unable to obtain valid student test scores for a given teacher since test scores within a content area include scores from students who were taught that content by a different teacher. For example, Teacher X teaches math and language to her homeroom, but her students go to Teacher Y for reading instruction. Standardized testing is done in the student's homeroom and all scores, including reading, are reported on the print-outs as if Teacher X gave the instruction even though they do not reflect Teacher X's instructional input.

Teachers who taught intact classes rather than regrouped classes were included in the sample. Only teachers of grades 1 - 6 were included as regrouping is nearly universal in the higher grades.

Grade by grade matching was done with an additional match within first grade classes for Chapter 1 and Transitional 1 classes. These classes are composed of scholastically disadvantaged first grade students. Chapter 1 and Transitional classes were identified within the CL sample and matched with Chapter 1 and Transitional classes within the NCL sample. Whenever possible, matches were obtained from within the same school. When this was not possible, CL teachers were matched with NCL teachers from a school with similar characteristics. Characteristics of the schools were determined by means of a composite School-score (Table 1) based on percentages of mobility, free lunch, 2 parent family, minority status and the Otis Lennon School Ability Index. For example, "A" School (50.92) was matched with "B" School (51.42).

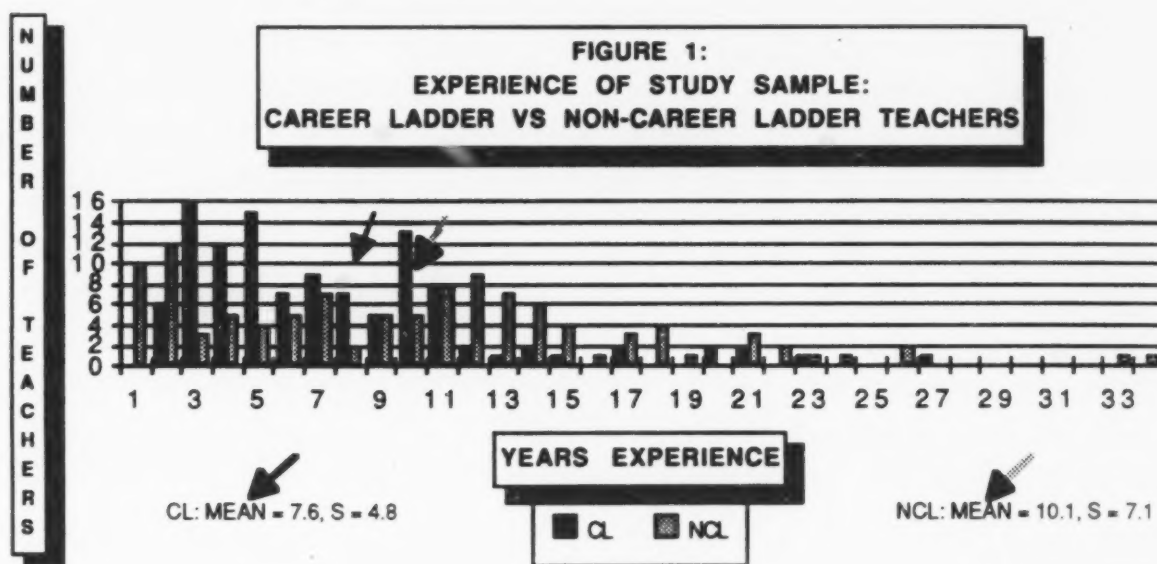
**Table 1: Means and Standard Deviations
of Matching Characteristics for Samples**

	School Scores		Experience		Class size	
	CL	NCL	CL	NCL	CL	NCL
Study 1	89.8 (27.4)	90.7 (27.2)	7.6 (4.8)	10.1 (7.1)	24.8 (4.5)	24.7 (4.4)
Study 2	100.6 (29.2)	101.6 (25.0)	8.88 (5.0)	9.6 (7.1)	25.6 (3.9)	26.0 (4.0)

Matches were also based on class size (Table 1). Where possible, a class of equal size within the same school was selected. If there was no class of equal size, the next closest size was selected. If there were several classes of the same size or equally close in size, random selections were made.

Because of the limited number of teachers, we were unable to match the samples on the basis of teacher experience although we initially believed this to be desirable. We collected self-reported data on total years experience for each teacher. The average years experience for non-career ladder teachers

was somewhat higher than for Career Ladder teachers (Table 1). Teacher experience for the first year sample is graphed in Figure 1.



It is apparent that experience differences exist between career ladder and non-career ladder teachers. Career ladder teachers cluster primarily between 2 and 11 years of experience while non-career ladder teachers are much more widely distributed with a bimodal distribution at 1-2 years and again between 9-15 years.

Year Two. Our second sample consisted of 244 teachers in grades 2-6 of whom 122 were placed on the District career ladder program and 122 were not. Of the second year non-career ladder sample, 9 teachers had been on CL during year one, but only data from 3 of these teachers had been included in the year 1 sample. First grade teachers were not included in the second sample because first grade students were excluded from the state testing. The 2 groups were again matched on the basis of class regrouping, school characteristics, grade and class size (Table 1) using the same procedures as were employed in Study 1. Information on teacher experience was again included for statistical control. However, the advantage in experience for

included for statistical control. However, the advantage in experience for non-career ladder teachers had now narrowed from 2.5 years difference to 0.8 years difference (Table 1).

MEASURES

Year One and Two. Testing was accomplished using the *Iowa Test of Basic Skills* (ITBS) administered district-wide in April of each year to all student in grades 1-6, the first year of our study and grades 2-6 the second year. The ITBS is a nationally standardized test developed to measure the educational progress of students in such content areas as reading, language, work-study and mathematics. Renorming of the ITBS occurred in the interim between year one and two. Where conversions were necessary, we used normative data from the Arizona Department of Education.

Construct validity of the ITBS was established with intercorrelation of scores between the ITBS and the *Test of Achievement and Proficiency* (TAP). Correlation ranges for each subtest are as follows:

Reading:	.69 - .93
Math:	.70 - .89
Language:	.67 - .91
Composite:	.77 - .87

Scores demonstrate stability over 1 to 3 year periods of time. Median correlations for 1-3 years by test are:

Reading:	.80 - .84
Math:	.81 - .83
Language:	.83 - .88
Composite:	.90 - .94

Reliability was reported using the Kuder-Richardson Formula 20 (KR-20) which is generally recognized as defining the lower limits of reliability. The KR-20 Reliability coefficients for grade 3 are representative of the general pattern across grades. These coefficients were reading (.91), math (.93),

language (.96) and composite (.98) indicating a high degree of reliability (Hieronymus, Hoover & Lindquist, 1986).

DESIGN

Year One. Two groups (career ladder teachers and non-career ladder teachers) were compared by grade (1,2,3,4-6) on average class Normal Curve Equivalent (NCE) scores from 4 sections of the Iowa Test of Basic Skills (math, verbal or reading comprehension, language and composite). Normal curve equivalents represent student achievement compared to a norm group. NCE scores were selected because they are more nearly comparable across grade levels and across tests than other scores and they can be averaged. Due to the regrouping problem we did not have adequate numbers of scores in grades 4-6 to analyze classes individually and these grades were subsequently combined and analyzed as a single grade designated "grade 4". Univariate analyses of covariance were accomplished by grade and by content area regressing on total years experience. A priori tests of parallelism demonstrated no significant interactions.

Year Two. The design for year two again used 2 groups (career ladder and non-career ladder) and compared by grade (2,3,4-6) on average class NCE scores from the same sub-sections of the ITBS as were used in study one. District policy changes in testing excluded first grade students and these scores were not available for year two, leaving us with 3 groups rather than the 4 we used in year one. Regrouping of classes again required us to combine grades 4-6. Univariate analysis of covariance by grade and content area regressing on education were analyzed at $\alpha = .05$. Tests of parallelism again were not significant.

ANALYSIS

We compared ITBS scores for all CL teachers with NCL teachers. Table 2 displays the adjusted means for these analyses.

Table 2
MEANS FOR IOWA TEST OF BASIC SKILLS
Career Ladder Status (Career vs Non-career) by
Grade Regressing on Experience (Year 1 & 2)

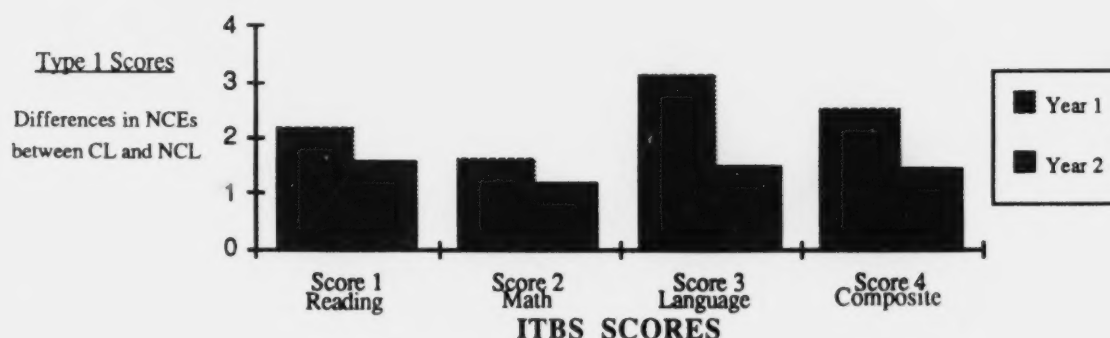
STATUS:	GRADE 1	GRADE 2	GRADE 3	GRADE 4-6
CAREER LADDER				
SCORE 1: READING				
Year 1:	50.89	60.12	55.11	56.27
Year 2:		54.79	51.61	53.65
SCORE 2: MATH				
Year 1:	54.45	61.16	55.97	55.62
Year 2:		58.14	49.38	53.97
SCORE 3: LANGUAGE				
Year 1:	61.23	68.18	63.22	60.84
Year 2:		59.89	53.77	54.42
SCORE 4: COMPOSITE				
Year 1:	56.01	64.46	59.17	58.31
Year 1:		59.28	52.13	54.69
NON-CAREER LADDER				
SCORE 1: READING				
Year 1:	48.51	58.21	52.57	54.36
Year 2:		51.49	50.68	53.22
SCORE 2: MATH				
Year 1:	52.49	60.59	53.55	54.16
Year 2:		56.63	47.99	53.36
SCORE 3: LANGUAGE				
Year 1:	58.28	65.06	59.55*	58.07
Year 2:		57.70	52.94	52.94
SCORE 4: COMPOSITE				
Year 1:	52.99	62.24	56.25	56.45
Year 2:		57.00	51.05	53.76

Year One. Classes taught by career ladder teachers consistently demonstrate higher means on all measures than do non-career ladder teachers (Table 2). Using a univariate approach, regressing on experience, we compared the career ladder group with the non-career ladder group. While means on all scores were consistently higher for CL teachers, only scores for the Grade 3

language section of the ITBS demonstrate statistical significance ($F(1,49) = 4.53$, $M_s = 46.50$). We then subtracted the scores of non-career ladder teachers from those of career ladder teachers to derive Type 1 scores demonstrating actual differences in student achievement (Figure 2).

Year Two. Again we found that classes taught by CL teachers produced higher means on all measures than do classes taught by NCL teachers (Table 2). These differences were not statistically significant. We subtracted the scores of non-career ladder teachers from those of career ladder teachers to derive Type 1 scores demonstrating differences in student achievement (Figure 2). Across grade and content areas, CL teachers' classes consistently outperformed those of NCL teachers an average of 1.41 NCEs (SD: .79) over the 2 year period.

Figure 2: Comparison of Achievement Score Differences (Type 1 Scores) by Year for Career Ladder and Non-Career Ladder Teachers.



Type 1 Score: Derived by subtracting the average achievement score of classes taught by non-career ladder teachers from that of career ladder teachers. Calculated from Normal Curve Equivalents which have a range of 0-99 and a mean of 50. This score demonstrates the degree to which career ladder teachers' student performance exceeded that of non-career ladder teachers.

The Career Ladder group was then segregated into higher Resident level and lower Career Level for comparisons. Adjusted means for these two groups of career ladder teachers are seen in Table 3.

TABLE 3
MEANS FOR IOWA TEST OF BASIC SKILLS
Career Ladder Status (Resident level vs career level)
By Grade, Regressing on Experience (Year 1 only)

CAREER LADDER GROUP	GRADE 1	GRADE 2	GRADE 3	GRADE 4
RESIDENT LEVEL				
SCORE 1: READING	43.51	59.72	53.31	54.95
SCORE 2: MATH	45.64	61.58	53.64	54.68
SCORE 3: LANGUAGE	53.26	66.70	61.74	60.58
SCORE 4: COMPOSITE	48.16	63.57	57.18	57.78
CAREER LEVEL				
SCORE 1: READING	54.95	60.36	57.21	57.16
SCORE 2: MATH	59.29	60.91	58.67	56.24
SCORE 3: LANGUAGE	65.61	69.07	64.96	61.01
SCORE 4: COMPOSITE	60.33	65.00	61.48	58.66

Year One. Means for the Career Level are consistently higher than means for the Resident Level. When we conducted ANCOVAs on these scores, again regressing on experience, we found no significant statistical differences between the two Career ladder levels.

Year Two. The number of teachers qualifying at the Resident Level of the career ladder was not sufficient to enable us to make comparisons between this group and either the Career Level of the ladder or the non-career ladder teachers.

Teachers at the higher Career Level only within the CL group were then compared with NCL teachers. Adjusted means for this comparison are seen in Table 4.

TABLE 4
ADJUSTED MEANS FOR IOWA TEST OF BASIC SKILLS:
Career Ladder Status (Career Level vs Non-career) by
Grade, Regressing on Experience (Years 1 & 2)

STATUS	GRADE 1	GRADE 2	GRADE 3	GRADE 4
CAREER LEVEL				
SCORE 1: READING				
Year 1:	54.95	60.36	57.21	57.16
Year 2:		55.48	52.41	54.38
SCORE 2: MATH				
Year 1:	59.29	60.91	58.67	56.24
Year 2:		59.43	50.01	55.78
SCORE 3: LANGUAGE				
Year 1:	65.61	69.07	64.96	61.01
Year 2:		61.18	53.76	55.67
SCORE 4: COMPOSITE				
Year 1:	60.33	65.00	61.48	58.66
Year 2:		60.51	52.67	55.90
NON-CAREER				
SCORE 1: READING				
Year 1:	48.51	58.21	52.57*	54.36
Year 2:		51.49	50.68	53.22
SCORE 2: MATH				
Year 1:	52.49	60.59	53.55	54.16
Year 2:		56.63	47.99	53.36
SCORE 3: LANGUAGE				
Year 1:	58.28	65.06	59.55*	58.07
Year 2:		57.70	52.94	52.94
SCORE 4: COMPOSITE				
Year 1:	52.99	62.24	56.25	56.45
Year 2:		57.00	51.05	53.76

* $p < .05$

Year One. Again we find the means for Career Level teachers consistently outpace those of NCL teachers. Significant differences were seen in Year 1 among Grade 3 scores in reading ($F(1,35) = 3.95$, $MS = 45.79$) and language (F

(1,35) = 4.85, MS = 52.37) Scores for the Grade 3 composite test and the Grade 1 language and composite tests approached significance.

Year Two. Across grade and content area, Career Level teachers' class means continued to be consistently higher than NCL teachers' class means. These differences did not prove to be statistically significant.

Year Two. Further isolating the highest levels of career ladder teachers (Career Level 5), we found only grade 2 with enough teachers who did not regroup to justify analysis. The means for Grade 2 by content area can be seen in Table 5. Analysis of variance revealed no significant differences.

TABLE 5 ADJUSTED MEANS FOR IOWA TEST OF BASIC SKILLS: GRADE 2 ONLY: Career Ladder Status (Career Level 5 vs Non-Career Ladder) by Grade, Regressing on Experience: (Year 2 Only.)				
	Score 1	Score 2	Score 3	Score 4
Career Level 5:	57.00	62.62	64.00	63.20
Non Career Ladder:	51.49	56.63	57.70	57.00

Year One. We had anticipated teacher experience might be an important variable in examining student achievement and subsequently regressed on that variable. When we examined this covariate, we found a significant main effect for experience in 9 of the 16 ANCOVAs we conducted. For overall career ladder versus non-career ladder teachers we found a main effect for scores 2 and 3 at grade 2. For Resident Level versus non-career we found no main effects. For Career Level versus non-career level we found a main effect for scores 2,3, and 4 at grade 2. For Career Level versus Resident Level we found a main effect for score 2 at grade 1 and for scores 2,3, and 4 at grade 2. No main effects were found for grades 3-6.

Year Two. Since study one, the differences in teacher experience narrowed substantially (Table 1) and we found no main effect for experience in any grade or content area in year two.

DISCUSSION

The Career Ladder Teacher Incentive Program was designed by the Arizona legislature to attract, motivate and retain talented teachers with both professional recognition and monetary incentives. If this program is functioning as designed, it should be selecting teachers whose students score higher on achievement testing. Using carefully matched samples of career ladder and non-career ladder teachers we compared mean class scores between and within these groups on the ITBS in a two year study.

When we subjected these means to analysis, the differences did not achieve statistical significance in most cases. Interpreting these findings in the strictest statistical sense would bring us to conclude that differences were incidental and that there is no benefit from the Career Ladder program nor is it functioning to select talented teachers. However, in the strictest statistical sense we would also anticipate patterns of variability in scores that we do not find here. Instead, we see a remarkably consistent pattern where CL teachers' scores were higher than NCL teachers' scores and higher for higher levels within the CL program itself.

If we examine mean score differences across scores and tests between Career Ladder teachers and non-career ladder teachers, we find an average difference of 2.35 points ($S=.75$) for year one and 1.41 ($S=.79$) for year two (Table 2) giving us an average difference across both years of 1.88. Isolating levels within the program itself, Career Level teachers scores averaged 4.68 points higher ($S=4.85$) than Resident Level teachers (Table 3) and Career Level teachers scores averaged 4.23 points ($S=2.14$) higher year one and 2.37 ($S=.94$) year two than non-career ladder teachers (Table 4). Even though differences were not statistically significant, the consistency of these patterns and the very fact that classes with Career Ladder teachers scored an average of from 2 to nearly 6 points higher than those taught by non-career ladder teachers makes us wonder if there are educational implications for this difference.

In an effort to appreciate the magnitude of these differences educationally, we engaged in a small game of logic. We know from our prior calculations

that across both years students of CL teachers averaged 1.88 points higher than students of NCL teachers on student achievement tests. If we look at the 2 year range of all classroom scores (47.99 to 68.18) and arbitrarily select the midrange score (58.08) as typical of NCL teachers, we can then assign a score 1.88 points higher as typical of CL teachers. Converting those NCE scores to percentiles (Hieronymous, Hoover and Lindquist, 1986), we can place NCL teachers' scores at the 64th percentile and CL teachers' scores at the 68th percentile. Class scores for CL teachers now demonstrate a 4 percentile advantage over those of NCL teachers. It becomes easier now to extract the educational meaning in these differences if we relate them to class size.

Glass and Smith (1978) proposed that as class size decreases, student achievement increases and demonstrated a 10 percentile rank achievement advantage for students taught in a class of 20 versus those taught in a class of 40. Our data demonstrated a 4 percentile achievement advantage for students taught by a career ladder teacher over those with NCL teachers. Placed in this context, the educational significance of even a few points becomes significant. The average class size in Mesa Public Schools is 26.7 students (Dickson, 1989). According to Glass and Smith's data, increasing student achievement by 6 percentiles in a class of 26.7 students places the achievement levels of these students near those which might be expected of a class with an average of 15 students.

Recall, however, that these data reflect only the first two years of implementation of the career ladder program at MPS. As with any new undertaking, there is a certain amount of backing and filling involved before the desired result is achieved. We believe it reasonable to assume that the results will become clearer as goals and objectives are clarified, procedures and reporting become better standardized and wider teacher acceptance and participation increase the scope of the program.

Workshops and study sessions designed to aid in application and student evaluation and to encourage career ladder teachers in meeting the program requirements were not instituted during the first year. For many teachers there was confusion in what was required and we believe the newness of the program served to temporarily focus participant energy on procedures of

enrollment rather than on upgrading teaching quality. Questions in later Career Ladder workshops often centered on how to meet the requirements of being accepted on Career Ladder Status and complete required paper work in a timely fashion rather than on demonstrating student achievement. As teachers are placed on the program this energy will be directed toward the classroom.

As we proposed previously, we believe a certain component of professional development is implicit in the career ladder program. Participation in on-going workshops and classes emphasizing student achievement increase the awareness of educators of the importance of this measure. Program requirements include evaluation of students and of teaching methods and this process alone should increase teacher effectiveness. In addition, collegiality is a viable part of the MPS career ladder design, and this collegiality encourages sharing and evaluation of various methods, contributing to the professional development that we believe is inherent in the program.

Because the Career Ladder Program at MPS is relatively new, there are some teachers who have not enrolled in it who undoubtedly would be accepted. Test scores of non-career ladder teacher may be influenced by the inclusion of scores of from this non-participant group and those of the CL group influenced by teachers who should not be there. This would tend to minimize differences. As selection procedures are refined and program acceptance increases, we hope to minimize this problem.

One variable one which we were unable to match our groups, was that of experience and as we have demonstrated, the matching group was somewhat more experienced than the Career Ladder group. To mitigate this effect, we regressed on experience, but found little evidence to indicate that experience is highly related to overall student achievement. The only grade where we were able to demonstrate a main effect for experience with any degree of consistency was with second grade students in math and language areas.

In summary, our findings demonstrate differences in student achievement, but those differences were not always statistically significant. The pattern of scores we found was Career Ladder/career level > Career Ladder/resident

level = Non-Career Ladder. Further examination and translation of these score differences into more meaningful terms provided better appreciation of the magnitude of education differences. Student scores in classes taught by career ladder teachers were 4 percentiles higher than scores of students taught by non-career ladder teachers. Given the intent of the legislation creating the pilot program was to differentially reward excellence in teaching, the program in Mesa Public Schools is achieving that goal.

References

- Cornett, L. (1987). More pay for teachers who do more: Incentive pay programs. 1987. Southern Regional Education Board Career Ladder Clearinghouse. Atlanta: pp 27.
- Dickson, L.S. (1989). Class size in Mesa Public Schools: Research and implementation. Department of Research and Evaluation, Mesa Public Schools, Mesa, AZ.
- Dickson, L.S. and Guy, V. (1990). Teacher attitudes toward performance-base pay. Paper presented at annual meeting of American Educational Research Association, Boston.
- ERS (Educational Research Services), (1978). Merit pay for teachers. ERS, Arlington, VA: , pg 126.
- ERS (Educational Research Service) (1983). Merit pay for teachers: Status and descriptions. ERS, Arlington, VA: pp186.
- Glass, G.V. and Smith, M.L. (1978). Meta-analysis of research on the relationship of class size and achievement. San Francisco: Far West Laboratory from Educational Research and Development.
- Hieronymus, A.N., Hoover, H.D. and Lindquist, E.F. (1986). Iowa Test of Basic Skills: National Norms for Forms G and H. Levels 5-14. Riverside Publishing Company, Chicago, page 107.
- Mesa Public Schools, (1986). Career ladder compensation plan: with revisions (12/86). V.L. Guy, Director, Career Ladder Compensation Plan, Mesa Public Schools, Mesa, AZ: pp 45.
- Mitchell, D.E. & Peters, M.J. (1988) A stronger profession through appropriate teacher incentives. Educational Leadership, 46, 3: pp 74-78.
- Packard, R.D. & Dereshiwsy, M.I. (1988) Study of the effects of a career ladder intervention program with focus on the production and outcomes in student achievement. Paper presented at the annual meeting of the Arizona Educational Research Organization, Phoenix, Nov. 1988.

Peterson, K.D. (1987). Teacher evaluation with multiple and variable lines of evidence. American Educational Research Journal, 24, 2, pp 311-317.

Robinson, G.E. (1984). Incentive pay for teachers: An analysis of approaches. *Concerns in Education*, March, 1984, Educational Research Service, Arlington, VA: pp 23.

USDE (U.S. Department of Education), (1983). A Nation at risk: The imperative for educational reform. The National Commission on Excellence in Education, T.H. Bell, Secretary of Education, Washington, D.C.: pp 64.

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991